OIPE

#7

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/821,839

DATE: 02/13/2002
TIME: 12:50:34

Input Set : A:\PSU0020.ST25.txt
Output Set: N:\CRF3\02132002\I821839.raw

ENTERED

```
 3 <110> APPLICANT: Ma, Hong
 5 <120> TITLE OF INVENTION: Plant Gene Required for Male Meiosis
 7 <130> FILE REFERENCE: Psu-0020
 9 <140> CURRENT APPLICATION NUMBER: 09/821,839
10 <141> CURRENT FILING DATE: 2001-03-29
12 <150> PRIOR APPLICATION NUMBER: 60/193,523
13 <151> PRIOR FILING DATE: 2000-03-31
15 <160> NUMBER OF SEQ ID NOS: 7
17 <170> SOFTWARE: PatentIn version 3.1
19 <210> SEQ ID NO: 1
20 <211> LENGTH: 2144
21 <212> TYPE: DNA
22 <213> ORGANISM: Arabidopsis thaliana
24 <400> SEQUENCE: 1
25 actgcatcag cccactctct agtctctgac taacgaactt ccattttcaa aattcgaatt    60
27 tctaatttct agtttcaagc tttcgtacgg agaaaaaatg aaggagatcg cgatgaggaa   120
29 ttcaaagcgc aagcctgagc cgacgccgtt cgccgggaag aagctccggt cgacgcgatt   180
31 acgccggaag agagcacaga tctctcccgt tcttgttcaa tcacctctct ggagcaaaca   240
33 aatcggagtc tctgctgctt ctgtcgattc ctgctccgat ttgctagctg atgacaacgt   300
35 ttcctgtggt tcgagcagag tcgagaagag ctcgaatccg aagaagactc taattgaaga   360
37 ggtagaagtt tctaaacctg gttataatgt gaaggagacg attggtgatt cgaaatttcg   420
39 aaggattacg aggtcttact ctaagctaca caaggagaag gagggagatg agatcgaagt   480
41 aagcgaatcg tcttgtgttg attcgaattc tggtgctgga ttaaggagat tgaatgtgaa   540
43 gggaaataaa attaacgaca acgatgagat ctctttctca cgatccgatg tgaccttcgc   600
45 cggacatgtc tccaacagcc ggagtttgaa tttcgaatcg gagaataagg agagcgacgt   660
47 cgtttctgtc atatctggag ttgagtactg ttccaagttc gggagcgtta ccggaggagc   720
49 tgataacgaa gaaattgaaa tctccaagcc gagcagcttc gtggaagctg attcctctct   780
51 tggatcggcc aaggaattga agccggagct tgagatagtc ggatgcgtct ctgatctcgc   840
53 ttgctctgag aaattctcgg aagaggtttc ggattctctc gatgatgagt catctgagca   900
55 acgttcagag atatattcac agtattccga cttcgattac tcggattaca ctccgtccat   960
57 cttcttcgac tctggcagcg aattctctga gaaatcttcc tctgattctc ctatttcaca  1020
59 ttctcgctct ctgtacctcc agttcaagga acagttctgt agatccacga ttcccaacga  1080
61 ttttggatct tcttgcgagg aagaaattca ctctgaattg ctaaggtttg atgatgagga  1140
63 ggtggaagag agctatctaa ggctgaggga aagagaaaga agtcatgcat atatgcggga  1200
65 ctgtgctaag gcatactgct ccaggatgga caatactggt ctcatccctc gtctacgctc  1260
67 catcatggtt caatggattg taaagcaatg ttctgacatg gggcttcagc aagagacatt  1320
69 gtttctagga gttggtctgt tggatcgatt cctgagcaaa ggatcattca aaagcgaaag  1380
71 gactctaata ctagtcggga ttgcgagtct tactctggcc accagaattg aagaaaatca  1440
73 accttacaac agcatccgga aaaggaactt caccattcag aacctaagat atagccggca  1500
75 tgaagtggtg gcaatggagt ggctggttca agaagtcctc aacttcaaat gcttcacacc  1560
77 cacaatcttc aacttcttgt ggttctactt aaaagctgct cgagccaatc cagaagttga  1620
79 aaggaaagcc aaatccttgg ctgttacctc actatccgac caaactcaac tctgtttttg  1680
81 gccctcaact gtagcagctg cactcgtggt tctcgcctgc atcgaacaca acaaaatctc  1740
```

2/13/02

```
83  tgcataccaa cgagtcataa aggtccatgt tagaacaaca gataacgagt tgcctgaatg   1800
85  cgttaagagt ctggactggt tgcttgggca gtaagcaatc aaaaagaaca aaaaccctaa   1860
87  aaccaggaca cagtatactc cgataccaac acacaggtta tcattactat ttacaaaaac   1920
89  aaacacaagg taagtaataa gaactcctct acagatttat atacttaatc gagctggact   1980
91  taattagctc ttagtatacc aattattagt gccaccattt gtgtcgctca tacacattta   2040
93  tttcttattt tccctaattc attagactct catattctta aaaagaatat ttccttgttt   2100
95  gaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa                     2144
98  <210> SEQ ID NO: 2
99  <211> LENGTH: 578
100 <212> TYPE: PRT
101 <213> ORGANISM: Arabidopsis thaliana.
103 <400> SEQUENCE: 2
105 Met Lys Glu Ile Ala Met Arg Asn Ser Lys Arg Lys Pro Glu Pro Thr
106 1               5                   10                  15
109 Pro Phe Ala Gly Lys Lys Leu Arg Ser Thr Arg Leu Arg Arg Lys Arg
110             20                  25                  30
113 Ala Gln Ile Ser Pro Val Leu Val Gln Ser Pro Leu Trp Ser Lys Gln
114         35                  40                  45
117 Ile Gly Val Ser Ala Ala Ser Val Asp Ser Cys Ser Asp Leu Leu Ala
118     50                  55                  60
121 Asp Asp Asn Val Ser Cys Gly Ser Ser Arg Val Glu Lys Ser Ser Asn
122 65                  70                  75                  80
125 Pro Lys Lys Thr Leu Ile Glu Glu Val Glu Val Ser Lys Pro Gly Tyr
126                 85                  90                  95
129 Asn Val Lys Glu Thr Ile Gly Asp Ser Lys Phe Arg Arg Ile Thr Arg
130             100                 105                 110
133 Ser Tyr Ser Lys Leu His Lys Glu Lys Glu Gly Asp Glu Ile Glu Val
134         115                 120                 125
137 Ser Glu Ser Ser Cys Val Asp Ser Asn Ser Gly Ala Gly Leu Arg Arg
138         130                 135                 140
141 Leu Asn Val Lys Gly Asn Lys Ile Asn Asp Asn Asp Glu Ile Ser Phe
142 145                 150                 155                 160
145 Ser Arg Ser Asp Val Thr Phe Ala Gly His Val Ser Asn Ser Arg Ser
146                 165                 170                 175
149 Leu Asn Phe Glu Ser Glu Asn Lys Glu Ser Asp Val Val Ser Val Ile
150             180                 185                 190
153 Ser Gly Val Glu Tyr Cys Ser Lys Phe Gly Ser Val Thr Gly Gly Ala
154         195                 200                 205
157 Asp Asn Glu Glu Ile Glu Ile Ser Lys Pro Ser Ser Phe Val Glu Ala
158     210                 215                 220
161 Asp Ser Ser Leu Gly Ser Ala Lys Glu Leu Lys Pro Glu Leu Glu Ile
162 225                 230                 235                 240
165 Val Gly Cys Val Ser Asp Leu Ala Cys Ser Glu Lys Phe Ser Glu Glu
166                 245                 250                 255
169 Val Ser Asp Ser Leu Asp Asp Glu Ser Ser Glu Gln Arg Ser Glu Ile
170             260                 265                 270
173 Tyr Ser Gln Tyr Ser Asp Phe Asp Tyr Ser Asp Tyr Thr Pro Ser Ile
174         275                 280                 285
177 Phe Phe Asp Ser Gly Ser Glu Phe Ser Glu Lys Ser Ser Ser Asp Ser
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/821,839

Input Set : A:\PSU0020.ST25.txt
Output Set: N:\CRF3\02132002\I821839.raw

```
178        290                   295                  300
181 Pro Ile Ser His Ser Arg Ser Leu Tyr Leu Gln Phe Lys Glu Gln Phe
182 305                   310                 315                  320
185 Cys Arg Ser Thr Ile Pro Asn Asp Phe Gly Ser Ser Cys Glu Glu Glu
186                   325                 330                  335
189 Ile His Ser Glu Leu Leu Arg Phe Asp Asp Glu Glu Val Glu Glu Ser
190                   340                 345                  350
193 Tyr Leu Arg Leu Arg Glu Arg Glu Arg Ser His Ala Tyr Met Arg Asp
194             355                 360                 365
197 Cys Ala Lys Ala Tyr Cys Ser Arg Met Asp Asn Thr Gly Leu Ile Pro
198     370                 375                 380
201 Arg Leu Arg Ser Ile Met Val Gln Trp Ile Val Lys Gln Cys Ser Asp
202 385                 390                 395                  400
205 Met Gly Leu Gln Gln Glu Thr Leu Phe Leu Gly Val Gly Leu Leu Asp
206                 405                 410                  415
209 Arg Phe Leu Ser Lys Gly Ser Phe Lys Ser Glu Arg Thr Leu Ile Leu
210             420                 425                 430
213 Val Gly Ile Ala Ser Leu Thr Leu Ala Thr Arg Ile Glu Glu Asn Gln
214             435                 440                 445
217 Pro Tyr Asn Ser Ile Arg Lys Arg Asn Phe Thr Ile Gln Asn Leu Arg
218     450                 455                 460
221 Tyr Ser Arg His Glu Val Val Ala Met Glu Trp Leu Val Gln Glu Val
222 465                 470                 475                  480
225 Leu Asn Phe Lys Cys Phe Thr Pro Thr Ile Phe Asn Phe Leu Trp Phe
226             485                 490                 495
229 Tyr Leu Lys Ala Ala Arg Ala Asn Pro Glu Val Glu Arg Lys Ala Lys
230             500                 505                 510
233 Ser Leu Ala Val Thr Ser Leu Ser Asp Gln Thr Gln Leu Cys Phe Trp
234         515                 520                 525
237 Pro Ser Thr Val Ala Ala Ala Leu Val Val Leu Ala Cys Ile Glu His
238     530                 535                 540
241 Asn Lys Ile Ser Ala Tyr Gln Arg Val Ile Lys Val His Val Arg Thr
242 545                 550                 555                  560
245 Thr Asp Asn Glu Leu Pro Glu Cys Val Lys Ser Leu Asp Trp Leu Leu
246             565                 570                 575
249 Gly Gln
253 <210> SEQ ID NO: 3
254 <211> LENGTH: 3018
255 <212> TYPE: DNA
256 <213> ORGANISM: Arabidopsis thaliana
258 <400> SEQUENCE: 3
259 gtcgaccaga gtttgaccaa tgactaatgt tatcgtatga ttcaattatt ttttgtacag    60
261 taatgtctcg tagaccgaca acaagaccag aaagtaatct taaaacccta gcttcacact   120
263 tagatatgtt tctatacatg tgtgtatata tacacacata cttgtagagc atgtgagaca   180
265 gtgaaattta ctagtgttaa acatggatgt gaagaaagat gagaacagta tcttagaaaa   240
267 catgaaacaa gagattaatc atagtctaaa agaagaagca caagaagaag aagagattct   300
269 aaagaagaga atctcaagcc accctttgta tgggcttctt cttcactcac atctcaattg   360
271 tttaaaggta cccctcttaa tgcttccctc tctctctttt ttattaaagt gatgtatgag   420
273 tataaatgtt tatctcttat gtatttggac ccaaccacga ccaggtgtgt tccggcgact   480
```

```
275 ttgactcacc ggagatcatg aacacggctg atgatcttgc cctatccaaa ctctctctcc   540
277 accctgactc ttcctccgaa gctacctctt cagaacttga tcaattcatg gttctttttt   600
279 ttttttctcc ctgccaaaac atatttacac aacaaaaaac aacttttcac gttcttttat   660
281 tttttccttt gcaaattaat ttgactttca agtattctaa gtttattttg ccaagaaaaa   720
283 aacagtagca cttatacttg taaattgatt cacataatag gaagcgtatt gctcgactct   780
285 acgggagctc aaggaagcaa tggagaagcc tcttaccgaa acgcatcgtt ttgtggatgc   840
287 ggtgtacact cagctaaacg acatcgttat gtcatcaccc ccttaaaaaa gggtaacatg   900
289 aacaactgtt cggtgctact atgtcaatgc attttgccaa attactactc agtctactca   960
291 cgatttattg tactgcgttt acgtaacgcg tttgtatgat cgtttattgg taaccgtaat  1020
293 ttatggcatg ccctcctgct tttttattta agaaaaataa aactaattat attgtaaata  1080
295 ttgcattgat catttagtca cactctttag aaaacaacag taaaatttaa atataaaaac  1140
297 aacactagct tccatgatta tttttcataa ccatttataa ttgcgtcatc ttgtaagttg  1200
299 taacgcattg cctttcttac tatgtaacgg ttgttgcata tttttgtgta cataaattta  1260
301 tacacaaaga taaaaagtga ctaagcttaa aatatccttg aaaaagcctt tgggtcatta  1320
303 acatggtgta agactacagg cgcattcagc aattggagtt ccgattctat tacagtaaga  1380
305 gggaacagaa ccgtaataat cgcgacacat ttgttcgcat ttgttagcat cgcatggaac  1440
307 cattggccag aaaacggggc aagtttgttc catcattctc gtctctctcg cacctttaaa  1500
309 caaacatcag aaaatttgtg acattaatta acaggatttg gcttcttata aagataagat  1560
311 taaaactact atttaaaaga taatctgtac ctgaggctga aacgatgaag atggtcatga  1620
313 taagaacagc gaaatttatg aggtttctca tggtttttatg tttttttttt tcttaacaaa  1680
315 gacgtaaact tgaatcgttt tatatgcgaa attgacagag aaaaccggaa aagataggat  1740
317 ctccttttct ttctttcttt tagtgaaata gatgataaac ttgtttctgc taaaagaggt  1800
319 gtttattttg gaaattatga attttctggt caatgtgatc ttagaatttt aaataggctg  1860
321 gattttgtga cctgattccg tgtcttatat ctgtatttac tatatttaga tgattctctg  1920
323 ataactgatg ttttaaaaag aagataattt tgataaagaa gtgattacga actttccaac  1980
325 attaaaagtt tagagtttat ttgattttat atctaatctt ggtttatatg tttttgatgg  2040
327 ggtttactaa ttatattata ccattcaagt tgaaatatat acaagttttt tttgtttat   2100
329 ccctaaattc tctaatgtga tatatataat atataatttg gatcggattc aaccaaacca  2160
331 tgaacgagat ttacattttg ccgttttccg aaatgttttg ggcttcgtaa agaactaaag  2220
333 gtgatattta gatattgggt atactatttg ttgtattggg cttaaaagtt tactttttttg  2280
335 gcccaaaatt aatcaactaa aataagatca ccaatggaaa aagaaacaaa aaaaccagta  2340
337 aaacatatgc agaaaatgta aatttacagg gcctaatata atctgcttga ccatgccatt  2400
339 gcgacataac aaatgttaca caagtagtgt acctataaag tagtgtacct ataatatatt  2460
341 aacagtgatc aatttcagtg tataaaaaaa gtcttcttaa atcatctttt aattccaaca  2520
343 atatgacatt cacaaactta tctatgattt ttttaaaaaa aaattcacac gtgtgctcaa  2580
345 tttatgtttc ttttagttct tccacgtgat ttgatgcaag aaaaatgatt agactgtatg  2640
347 ttaaaaagca tactagagaa attaattata aaacatcaat cagttgaagt aattatcaaa  2700
349 accgcatgct tttttagcta aatctgtgat tgtactgacg cagatgcata aattcaaacg  2760
351 caaacgctga tctctacatt agccaaacaa gaatagcgtc caaatttacg actggtttca  2820
353 cgtgcaccaa accgtagggt ataatatctc tctctcactc tccaacatcc ccactcttcc  2880
355 caagaaactt ctataactgc atcagcccac tctctagtct ctgactaacg aacttccatt  2940
357 ttcaaaattc gaatttctaa tttctagttt caagctttcg tacggagaaa aaatgaagga  3000
359 gatcgcgatg aggaattc                                                  3018
362 <210> SEQ ID NO: 4
363 <211> LENGTH: 3970
364 <212> TYPE: DNA
365 <213> ORGANISM: Arabidopsis thaliana
367 <400> SEQUENCE: 4
368 ttgaccatgc cattgcgaca taacaaatgt tacacaagta gtgtacctat aaagtagtgt    60
```

```
370 acctataata tattaacagt gatcaatttc agtgtataaa aaaagtcttc ttaaatcatc  120
372 ttttaattcc aacaatatga cattcacaaa cttatctatg attttttttaa aaaaaaattc  180
374 acacgtgtgc tcaatttatg tttctttttag ttcttccacg tgatttgatg caagaaaaat  240
376 gattagactg tatgttaaaa agcatactag agaaattaat tataaaacat caatcagttg  300
378 aagtaattat caaaaccgca tgcttttta gctaaatctg tgattgtact gacgcagatg  360
380 cataaattca aacgcaaacg ctgatctcta cattagccaa acaagaatag cgtccaaatt  420
382 tacgactggt ttcacgtgca ccaaaccgta gggtataata tctctctctc actctccaac  480
384 atccccactc ttcccaagaa acttctataa ctgcatcagc ccactctcta gtctctgact  540
386 aacgaacttc cattttcaaa attcgaattt ctaatttcta gtttcaagct ttcgtacgga  600
388 gaaaaaatga aggagatcgc gatgaggaat tcaaagcgca agcctgagcc gacgccgttc  660
390 gccgggaaga agctccggtc gacgcgatta cgccggaaga gagcacagat ctctcccgtt  720
392 cttgttcaat cacctctctg gagcaaacaa atcggagtct ctgctgcttc tgtcgattcc  780
394 tgctccgatt tgctagctga tgacaacgtt tcctgtggtt cgagcagagt cgagaagagc  840
396 tcgaatccga agaagactct aattgaagag gtagaagttt ctaaacctgg ttataatgtg  900
398 aaggagacga ttggtgattc gaaatttcga aggattacga ggtcttactc taagctacac  960
400 aaggagaagg agggagatga gatcgaagta agcgaatcgt cttgtgttga ttcgaattct 1020
402 ggtgctggat taaggagatt gaatgtgaag ggaaataaaa ttaacgacaa cgatgagatc 1080
404 tctttctcac gatccgatgt gaccttcgcc ggacatgtct ccaacagccg gagtttgaat 1140
406 ttcgaatcgg agaataagga gagcgacgtc gtttctgtca tatctggagt tgagtactgt 1200
408 tccaagttcg ggagcgttac cggaggagct gataacgaag aaattgaaat ctccaagccg 1260
410 agcagcttcg tggaagctga ttcctctctt ggatcggcca aggaattgaa gccggagctt 1320
412 gagatagtcg gatgcgtctc tgatctcgct tgctctgaga aattctcgga agaggtttcg 1380
414 gattctctcg atgatgagtc atctgagcaa cgttcagaga tatattcaca gtattccgac 1440
416 ttcgattact cggattacac tccgtccatc ttcttcgact ctggcagcga attctctgag 1500
418 aaatcttcct ctgattctcc tatttcacat tctcgctctc tgtacctcca gttcaaggaa 1560
420 cagttctgta gatccacgat tcccaacgat tttggatctt cttgcgagga agaaattcac 1620
422 tctgaagtaa gtggtataat gatttcatat ctcttggaat aattgctagt ggttagagat 1680
424 tgaagatgta tgtggttata tggttgaaat ttcattcgat tactagtcta tttttgatat 1740
426 gagacttgtt ctgctctgtg tttgattctg aaattttgtt ctggaatgaa tcttaagtat 1800
428 acattttcgt tttagttgct aaggtttgat gatgaggagg tggaagagag ctatctaagg 1860
430 ctgagggaaa gagaaagaag tcatgcatat atgcgggact gtgctaaggc atactgctcc 1920
432 aggatggaca atactggtct catccctcgt ctacgctcca tcatggttca atggattgta 1980
434 aaggtgaatt ttaactttct gttcaaatgc atttagttac atatacattg atctctgaat 2040
436 gttgaagctc agaaatatgt atcagtagca gaagattatg aagtaaatga atatttggag 2100
438 atcctgttcc tggtttttaag aatgtttttag cctaaggaaa tctatagctt actttggaat 2160
440 cttttaaggt ttatgtatca gtcagctatg atattctttg ttgctgattg tctgctccct 2220
442 gattacaagc agcaatgttc tgacatgggg cttcagcaag agacattgtt tctaggagtt 2280
444 ggtctgttgg atcgattcct gagcaaagga tcattcaaaa gcgaaaggac tctaatacta 2340
446 gtcgggattg cgagtcttac tctggccacc agaattgaäg aaaatcaacc ttacaacagg 2400
448 taccaaccat attccatctt catgattctg acttccaatg ttcattagaa aagtgttctg 2460
450 agtaggaaaa agattaggac cattacaaga aactgagtat tacgcttaac caaatcaagg 2520
452 actaataatg gtctaataca aacccttatg gttcaatgaa ttggcatttc atgtgggtat 2580
454 cgaatattgg attatgtttc tcaaaaacac tctttactgg aaagaacctt ccacaataca 2640
456 caggaatagt tcattttct tcaactgctc acctgatact tgctcttttt aactagcatc 2700
458 cggaaaagga acttcaccat tcagaaccta agatatagcc ggcatgaagt ggtggcaatg 2760
460 gagtggctgg ttcaagaagt cctcaacttc aaatgcttca cacccacaat cttcaacttc 2820
462 ttgtggtaaa acctctctga ctatatattt tcatgttcca agacacatta tccacacaga 2880
464 aagatacata tgactatcat ttatacatgt caggttctac ttaaaagctg ctcgagccaa 2940
466 tccagaagtt gaaaggaaag ccaaatcctt ggctgttacc tcactatccg accaaactca 3000
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/821,839

DATE: 02/13/2002
TIME: 12:50:36

Input Set : **A:\PSU0020.ST25.txt**
Output Set: **N:\CRF3\02132002\I821839.raw**